EFFICIENT RECOMMENDER SYSTEMS

By

Dirk Bergemann and Deran Ozmen

June 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1568

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Efficient Recommender Systems

Dirk Bergemann
Yale University
New Haven, CT 06511
dirk.bergemann@yale.edu

Deran Ozmen
Yale University
New Haven, CT 06511
deran.ozmen@yale.edu

April 2006
Extended Abstract

Abstract

We study the efficient allocation of buyers in the presence of recommender systems. A recommender system affects the market in two ways: (i) it creates value by reducing product uncertainty for the customers and hence (ii) its recommendations can be offered as add-ons, which generates informational externalities. We investigate the impact of these factors on the efficient allocation of buyers across different products.

We find that the efficient allocation requires that the seller with the recommender system has full market share. If the recommender system is sufficiently effective in reducing uncertainty, it is optimal to have some products to be purchased by a larger group of people than others. The large group consists of customers with flexible tastes.

1 Introduction

The large volume of transactions on the internet gives rise to a large accumulation of data about customers and products. This enables internet sellers to build databases that consist of personalized data on all their customers, the customers' past purchases and the feedback from those purchases. In this paper we analyze one specific use of the accumulated information, "recommender systems". A recommender system is a software program which uses the accumulated data to make statistical inferences about what product a particular customer would like when she returns to the website.

From an economic point of view, a recommender system represents an informational linkage that creates additional surplus by reducing uncertainty for the customers. In this paper we present a two-period, two-product model that describes the interaction between a seller employing a simple recommender system and a competitive fringe with no such system, to analyze the surplus created by recommender system. This paper takes an efficiency point of view and investigates how that surplus could be maximized.

There are usually two sources of uncertainty involved in the decision process of a customer. She may be unsure about her tastes and/or characteristics of the products. In our model, we focus only on product uncertainty in the on-line market for horizontally differentiated products, where the difference in customers' tastes translate into differences in the willingness to pay for decreased uncertainty. Our recommender system acts as a mechanism that collects customer evaluations, through which the seller infers more information about the products. The seller reveals whatever inference he makes to his "loyal" customers, those who have made a purchase from him before. Thus, a loyal customer has the chance to make a more informed choice using the inference revealed to her by the recommender system.

As we mentioned above, a recommendation can be considered as an add-on: it is an additional service a customer receives on top of the purchase she makes. Recommendations, however, are different from typical add-ons and bundle elements because their quality is determined endogenously by the information accumulated through the seller's sales. Thus the efficiency problem of what market share each seller should have entails informational externalities. These externalities can be separated into two elements. The first element is what we call the "volume externality". This externality represents the general coordination element inherent in the problem, which is that as the seller has more customers, he will be able to make better recommendations. This element determines how much of the market the seller should capture to maximize the surplus. The second one is the "product externality". This externality relates to the distribution of buyers within one seller over different products. If there are a lot of customers buying one particular product in one period, it might be worth having other customers delay the purchase of that product and purchase the other products for that period. The strength of this effect determines whether equal amounts of information should be accumulated on each product or whether there are increasing returns to information so that a large volume of buyers should be induced to buy some products and provide information at the expense of other products on which smaller volume of information is gathered. In the model, the customers differ both in the type of product they prefer and in the intensity of their preferences. Some buyers are more flexible in their choices than others. It is the buyers with inflexible tastes who really benefit from the recommendation service.

Avery, Resnick and Zeckhauser (1999) consider a recommender system with a single product and sequential choice. Ansari, Essegaier and Kohli (2000) describe and compare methods of prediction which range from classic linear regression to Bayesian methods.

2 Model

Market The supply side of the market consists of a seller with a recommender system, denoted by M, and a competitive fringe of sellers with no recommender system, denoted by F. The demand side consists of a continuum of buyers, and each buyer has a choice between one of two products. Each buyer is characterized by his preference parameter θ which is distributed uniformly over $[-1,1]$. The products are denoted by $x \in \{-1,1\}$. The gross utility of a buyer of type θ is specified by:

$$u(\theta,x) = v - (\theta - x)^2 . \tag{1}$$

As an example consider the product line to be books. Then the two types of the product can represent "mystery" versus "romance" novels. We can consider the buyers with preference parameters close to -1 or 1 as "inflexible" and buyers with preference parameter close to 0 as "flexible", because the former group would insist on their favorite kind of book whereas the latter group would not be adverse to trying other kinds.

Timing and Choices In each period new products arrive and there is some uncertainty about their characteristics. The sellers and buyers share a common prior on these products' types. In period 0 two products, denoted by l and h, arrive at both sellers. These products are differentiated with respect to the priors attached to them. Let $x_i \in \{-1,1\}$ be the true type of product $i \in \{l,h\}$ and $\alpha_i \equiv \Pr(x_i = 1)$. The two products arrive with symmetric uncertainty:

$$\begin{aligned} \alpha_h &= \tfrac{1}{2} + \varepsilon \\ \alpha_l &= \tfrac{1}{2} - \varepsilon \end{aligned} \tag{2}$$

where $\varepsilon \in \left[0, \frac{1}{2}\right]$. The initial priors are differentiated by ε, which we refer to as "initial information".

In period 1 a new product, m, arrives with prior $\alpha_m \in \{\alpha_l, \alpha_h\}$ at all sellers. In period 0, neither the buyers nor the sellers know what the exact value of α_m will be in period 1, but they attach $\frac{1}{2}$ probability to α_m being α_h and α_l. The time-line is presented below.

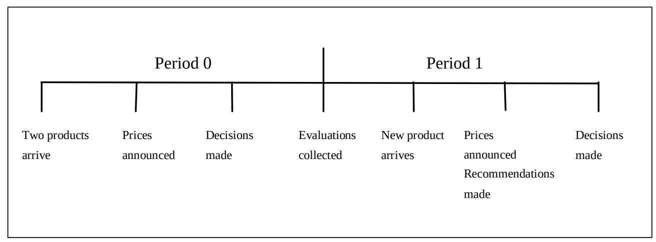

Figure 1: Timeline

Learning Between periods 0 and 1 seller M receives information form his buyers. We aggregate the information as follows: Let μ_i denote the measure of buyers who buy product $i \in \{l,h\}$ from seller M in period 0. Seller M receives a random signal $y_i(x_i) \in \{-1,0,1\}$ on the type of each product $i \in \{l,h\}$ between periods 0 and 1, where

$$\begin{aligned} \Pr(y_i(x_i) = 0 \mid x_i) &= 1 - \mu_i \\ \Pr(y_i(x_i) \in \{-1,1\} \mid x_i) &= \mu_i \end{aligned}$$

We can interpret a signal of 0 as containing no information, or simply the failure to receive an informative signal. Given that the seller receives a relevant signal, the probability of the signal being correct is:

$$\Pr(y_i(x_i) = x_i \mid y_i(x_i) \in \{-1,1\}, x_i) = \frac{1}{2} + \gamma,$$

where $\gamma \in \left[0, \frac{1}{2}\right]$. We interpret γ as the informativeness of the signal. Given the probabilistic structure, we view the recommender system as a mechanism that computes the posterior beliefs for each product i based on the signal

y_i and reports them only to the buyers who have bought from him in period 0. The posterior for product i given signal y_i will be denoted by

$$\alpha_i(y_i) \equiv \Pr(x_i = 1 \mid y_i).$$

Interpretation There are two products arriving with symmetric uncertainty attached in period 0. A high ε means there is less uncertainty about each product's type and that the two products are highly differentiated. A low ε means uncertainty is high for both products and that initially the two products look similar. On the other hand, γ represents the informativeness of the signal. We interpret ρ:

$$\rho = \frac{\gamma}{\varepsilon}, \tag{3}$$

as the "performance" of the recommender system.

3 Social Efficiency

In this section we analyze the efficient assignment problem of how to distribute buyers over different sellers and products in order to maximize total surplus.

In the absence of a recommender system, the efficient allocation in period 0 is straightforward. Each buyer should be allocated to the product that gives her the highest per period utility, i.e. all buyers with $\theta \geqslant 0$ should buy product h and all buyers with $\theta < 0$ should buy version l. Notice that the seller choice does not matter in this case, because there is no difference between the service provided by different sellers.

If we introduce uncertainty and information, it is no longer true that each buyer should buy the product which gives her the highest per period expected utility, because a buyer's choice of seller and product in period 0 affects the utility of all the other buyers in period 1. There are two variables which influence the informational externality: (i) the distribution of buyers over sellers M and F and (ii) the distribution of buyers over the two different products. The first variable is important because it determines the aggregate information gathered by seller M and thus the overall effectiveness of the recommender system in reducing uncertainty. It is clear that all buyers should purchase from seller M in period 0 because information has positive value and the inflow of information is maximized when seller M has full market share. Therefore, we ignore the fringe in the remaining analysis.

Definition 1 (BALANCE)
A distribution of buyers with (μ_h, μ_l) is balanced if $\mu_h = \mu_l$ and unbalanced if $\mu_i > \mu_j$ for some $i \in \{l, h\}$.

We would like to investigate whether it can be optimal to create endogenous differentiation through unbalanced distributions. If the distribution is unbalanced, one product is experimented by a larger group of buyers and the small group of buyers wait to benefit from their feedback. If this is the case, then it is also important for efficiency to know the composition of these groups. This suggests the following definition.

Definition 2 (SORTING)
A distribution of buyers (μ_l, μ_h) is:

1. *"sorted" if the set of buyers selecting products l and h are line segments of the form $[-1, x]$ and $[x, 1]$;*
2. *"shuffled " if $\mu_i \geqslant \mu_j$ for some $i \in \{l, h\}$ and the set of buyers selecting product j consists of two segments S^-, S^+ of the forms $[-1, x]$, $[y, 1]$. The degree of shuffling is:*
$$\frac{\min\{|S^-|, |S^+|\}}{\max\{|S^-|, |S^+|\}};$$
3. *"perfectly shuffled" if $|S^-| = |S^+|$.*

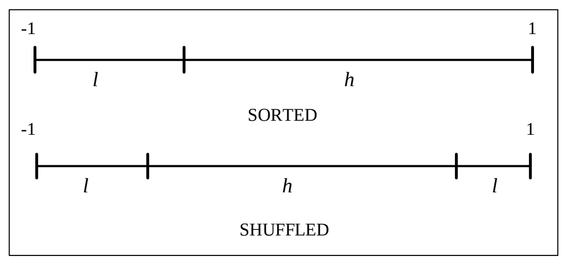

Figure 2: Sorting and Shuffling for $\mu_h > \mu_l$

If a distribution is shuffled, it is the inflexible buyers of both types that benefit more from the endogenous differentiation created by the unbalanced distribution. On the other hand, if a distribution is sorted, it is usually the inflexible buyers of one type receiving information from the experiences of all other buyers. The individual preference of a buyer becomes increasingly influential in the determination of the social surplus as the buyer's type moves further away from 0. Therefore, knowing the preference ranking of each buyer over the two products is essential to finding the right balance between the individual buyers' interests and the society's interest, which determines the efficient solution.

In the analysis of the efficient distribution (μ_l, μ_h), we have to be aware of the fact that the ranking of the utilities from the two products might change as we change

(μ_l, μ_h). The preference ranking between h and l depends on the difference between $|\mu_h - \mu_l|$. For example, if μ_h is sufficiently greater than μ_l, then even the buyers who would prefer product h in a static world, i.e. the buyers with $\theta > 0$, get a higher two-period utility from product l than product h.

It is also clear that once we restrict attention to full market share distributions, the only balanced distribution is $\left(\frac{1}{2}, \frac{1}{2}\right)$, where the rankings are in accordance with the static preferences. However, there could be different preference rankings over h and l when the distribution is unbalanced with $\mu_l < \mu_h$. Next we consider optimal shuffled distributions.

Lemma 1 *If a shuffled distribution with (μ_h, μ_l) is efficient, then $|S^-| \lesseqgtr |S^+|$ as $\mu_h \lesseqgtr \mu_l$.*

In other words, if it is efficient for some positive types to purchase product l, then it has to be the case that there are more negative types purchasing l.

Proposition 1 (EFFICIENCY)
In the efficient allocation seller M has a full market share and there is a unique $\rho_s > 3$ such that:

1. *for $\rho \leqslant \rho_s$, the unique efficient distribution of buyers is balanced and sorted;*

2. *for $\rho_s < \rho < \infty$, there are two symmetric efficient distributions of buyers that are unbalanced and imperfectly shuffled.*

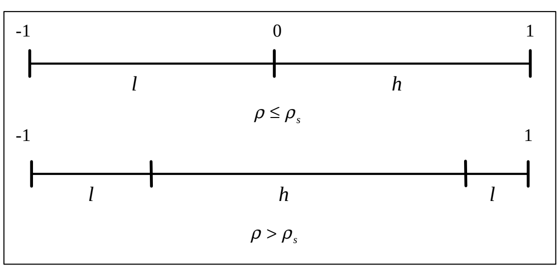

Figure 3: Efficient allocations with $\mu_h \geqslant \mu_l$

In other words, for low levels of ρ, the efficient allocation with a recommender system does not differ from the efficient allocation without a recommender system. With an increase in ρ, the information provided by the recommender system becomes sufficiently valuable to make the efficient distribution unbalanced.

Proposition 2 (COMPARATIVE STATICS)

1. *At $\rho = \rho_s$, the degree of unbalance and the degree of shuffling increase discontinuously.*

2. *For $\rho \geqslant \rho_s$, the degree of unbalance and shuffling increase in ρ.*

3. *As $\rho \to \infty$, the distribution for both efficient allocations becomes perfectly shuffled.*

As information becomes more valuable it is beneficial to increase the degree of unbalance and place a higher burden on flexible buyers. As there is more that the recommender system can contribute, the inflexible buyers increasingly have more to gain than flexible buyers.

4 Conclusion

We have shown how the existence of a recommender system creates additional surplus and introduces informational externalities into the problem of optimally distributing the buyers over different products. If the recommender system's output were independent of the sales, then buyers would be optimally allocated to the products in line with their static preferences. As the information produced by the recommender system is endogenously determined through the sales, it is not always optimal to allocate the buyers according to their static preferences. Typically, the future gains of the inflexible buyers more than offset the current losses of the flexible buyers. As the information produced by the recommender system becomes more valuable, more flexible buyers will acquire information to improve the future consumption choice of the inflexible buyers. Finally. in related work, Bergemann and Ozmen (2006) analyze how a profit maximizing seller can use a recommender system to improve its revenue management.

References

[1] Ansari, A.,S. Essegaier, and R. Kohli (2000): "Internet Recommendation Systems", *Journal of Marketing Research,* 37, 363-375.

[2] Avery, C., P. Resnick, and R. Zeckhauser (1999): "The Market for Evaluations", *American Economic Review*, 89, 564-584.

[3] Bergemann, D. and D. Ozmen (2006): "Optimal Pricing Policy with Recommender Systems", *ACM-EC 2006.*